

## Sample Examination Three

### Section I

Time—1 hour and 30 minutes

Questions 1–40

Percent of total grade—50

**Directions:** The questions or incomplete statements that follow are each followed by five suggested answers or completions. Choose the response that best answers the question or completes the statement.

1. If  $X$  is a random variable, which of the following does NOT represent a probability distribution of  $X$ ?

(A)

$X$	0	1	2	3
$P(X)$	.1	.35	.40	.15

(B)

$X$	-2	-1	0	1	2
$P(X)$	.13	.28	.18	.28	.13

(C)

$X$	1	2	3	4	5	6
$P(X)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

(D)

$X$	0	1	4	9
$P(X)$	.1	.2	.3	.4

(E)

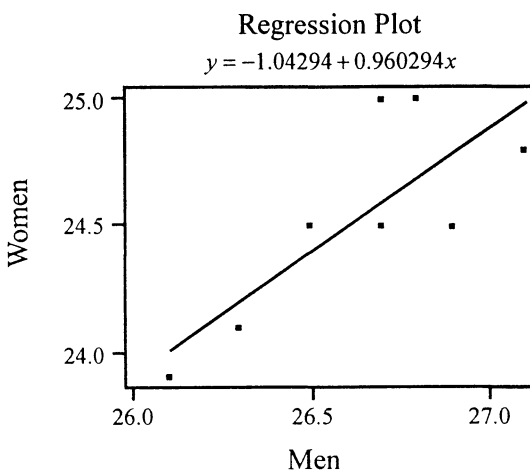
$X$	1	3	5	7	9
$P(X)$	.1	.2	.3	.4	.5

Answer

2. There is a linear relationship between the duration  $x$  (in seconds) of an eruption of a geyser and the interval of time  $y$  (in minutes) until the next eruption. A least-squares regression line of data collected by a geologist is represented by,  $\hat{y} = 41.9 + 0.18x$ ,  $100 < x < 300$ . What is the estimated increase in the interval of time until the next eruption that corresponds to an increase of 60 seconds in the duration?
- (A) 0.18 minutes  
 (B) 3.6 minutes  
 (C) 10.8 minutes  
 (D) 36.0 minutes  
 (E) 41.9 minutes

Answer

3. Using computer software and data for the years 1990–1998, a least squares regression line was computed to predict the median age of women  $y$  given the median age of men ( $x$ ), at first marriage.



The regression equation is

$$\text{Women} = -1.04 + 0.960 \text{ Men}$$

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Men	9	26.622	26.700	26.622	0.307	0.102
Women	9	24.522	24.500	24.522	0.373	0.124

The percent variation of  $y$  that is explained by the Least Squares Regression of  $y$  on  $x$  is

- (A) 0.04  
 (B) 0.08  
 (C) 0.62  
 (D) 0.79  
 (E) 0.96

Answer

4. A survey reveals that 60% of patients who receive prescription medications from their doctors prefer the brand-name drug as opposed to the generic brand. In a random sample of 40 patients, what is the probability that at least 25 of them prefer the brand name?

(A) 0.1255  
(B) 0.3745  
(C) 0.4402  
(D) 0.6250  
(E) 0.9500

Answer

5. A consequence of the Central Limit Theorem is that for  $n$  sufficiently large ( $n \geq 30$ ), if all samples of size  $n$  are taken, the mean of the sample means  $\mu_{\bar{x}}$  is equal to the population mean  $\mu$ . Since the mean of the sampling distribution is equal to the population mean,  $\bar{x}$  is referred to as

(A) a biased estimator  
(B) an unbiased estimator  
(C) a random estimator  
(D) a controlled variable  
(E) a parameter

Answer

6. A survey indicates that approximately 40% of college students own a cellular phone. In a random sample of 15 college students, what is the probability that at least 5 own a cellular phone?

(A) 0.010  
(B) 0.186  
(C) 0.217  
(D) 0.333  
(E) 0.783

Answer

7. In a family of seven children, what is the probability that at least one of the children is a girl? Assume the probability of a child being a girl is 50%.

(A)  $1/2$   
(B)  $1/7$   
(C)  $1/14$   
(D)  $1/128$   
(E)  $127/128$

Answer

8. The probability distribution tables for two random variables  $x$  and  $y$  are given below:

$x$	-2	-1	0	1	2
$P(x)$		.1	.3	.1	

$y$	0	1	2
$P(y)$	.1		.4

Suppose  $x$  and  $y$  are independent and  $P(x = 2, y = 1) = 0.2$ . What is  $P(x = -2)$ ?

(A) 0.1  
(B) 0.2  
(C) 0.3  
(D) 0.4  
(E) 0.5

Answer

9. A survey was conducted among 340 high school girls in which they were polled regarding whether they were right-handed or left-handed and whether or not they played a musical instrument. The table below displays the results of the survey:

	Play	Does not play	Total
Right-handed	74	236	310
Left-handed	18	12	30
Total	92	248	340

If the likelihood of a girl playing an instrument is independent of left or right-handedness, what is the expected number in the cell representing right-handed girls who play a musical instrument?

- (A) 62
- (B) 74
- (C) 80.2
- (D) 83.9
- (E) 92.1

Answer

10. A health fitness research group wishes to estimate the mean amount of time (in hours) that members of a fitness center spend each week exercising at the center. They want to estimate the mean within a margin of error of 0.5 hours with a 95% level of confidence. Previous data suggests that  $\sigma = 2.2$ . Which of the following is the smallest sample size that meets these criteria?

- (A) 60
- (B) 75
- (C) 90
- (D) 180
- (E) 190

Answer

11. A die has six faces, all of which are equally likely to occur. Three of the faces have the number 6 printed on them, two of the faces have the number 12 printed on them, and one face has a 0 printed on it. What is the expected value of the sum if one were to roll two of these dice?

- (A)  $\frac{11}{3}$   
(B) 12  
(C) 14  
(D) 18  
(E) 24

Answer

12. Which of the following associations between the two given variables is likely to have a negative correlation?

- (A) Number of hours devoted to studying for a final exam and a student's grade on the final  
(B) A teacher's salary and the number of years of teaching experience that the teacher has  
(C) The age of an automobile and the number of miles an automobile has  
(D) The number of children in a family and the weekly family food budget  
(E) The time and speed required to travel a given distance

Answer

13. Suppose that for a certain couple, the probability of a male birth is 0.47 and a female birth is 0.53. If the couple decides to continue having children until they have one son, what is the probability that they will have four children?

- (A) 0.05  
(B) 0.07  
(C) 0.19  
(D) 0.25  
(E) 0.47

Answer

14. A South Florida newspaper reports that the average cost of renting a car in Fort Lauderdale during the months of January and February is \$36 per day. A random sample was taken of 40 people who rented cars at Fort Lauderdale Airport and the mean of this sample was found to be  $\bar{x} = \$34.20$ . To test the claim that the average rental in Fort Lauderdale is less than that reported by the newspaper which of the following statements can represent the null hypothesis?

- (A)  $H_0: \mu < 36$
- (B)  $H_0: \mu \neq 36$
- (C)  $H_0: \mu = 36$
- (D)  $H_0: \mu < 34.2$
- (E)  $H_0: \mu = 34.2$

Answer

15. Data was collected to test whether there is a difference between the mean heights of women born in two different countries. It is assumed that the heights of women in each of these countries are normally distributed. Also, there is no reason to believe that population variances of the population of heights of women from the two different countries are equal. The data collected indicate the following:

$n_1 = 130$	$\bar{x}_1 = 63.8$	$s_1 = 2.3$
$n_2 = 190$	$\bar{x}_2 = 62.4$	$s_2 = 2.7$

Which of the following represents an estimate of the standard deviation of the sampling distribution corresponding to  $\bar{x}_1 - \bar{x}_2$ ?

- (A) 0.04
- (B) 0.29
- (C) 2.24
- (D) 2.50
- (E) 3.16

Answer

16. Two events A and B have the following probabilities:  $P(A) = 0.54$ ,  $P(B) = 0.20$ , and  $P(A \text{ and } B) = 0.108$ . Which of the following conclusions can be drawn from the data?
- (A) A and B are mutually exclusive events
  - (B) A and B are independent events
  - (C) A and B are dependent events
  - (D) A and B are complementary events
  - (E) Not enough information is given to draw a conclusion

Answer

17. In order to test reading abilities of fifth grade children, a representative sample of 120,000 fifth graders is to be selected in a heavily populated state. It is known that 60% of the children in that state live in urban areas, 25% live in suburban areas, and 15% live in rural areas. It is also known that reading ability levels are different in these three types of areas. Which of the following is the most appropriate sampling procedure to use?
- I. Draw a random sample of 40,000 in each of the three categories.
  - II. Draw random samples of 72,000 urban children, 30,000 suburban children, and 18,000 rural children.
  - III. Perform a random sampling of the schools, including enough classes so as to have 120,000 children in the sample.
- (A) I only
  - (B) II only
  - (C) III only
  - (D) I and II are both equally appropriate
  - (E) II and III are both equally appropriate

Answer

18. For a given hypothesis test, the probability of a Type I error is represented by  $\alpha$  and the probability of a Type II error is represented by  $\beta$ . Which of the following represents the power of the test?
- (A)  $\alpha$
  - (B)  $1 - \alpha$
  - (C)  $\beta$
  - (D)  $1 - \beta$
  - (E)  $\alpha + \beta$

Answer



19. The admissions policy at a certain university requires that incoming students score in the upper 20% on a standardized test. If the mean score on the test is 510 and the standard deviation of the scores is 80, what is the minimum score that a student can earn on the test to meet the admissions requirement? Scores on the test are normally distributed and are reported in intervals of 10.

(A) 520  
(B) 580  
(C) 590  
(D) 600  
(E) 620

Answer

20. In a sample survey taken of 450 residents of a given community, 180 of them indicated that they shop at the local mall at least once per month. Construct a 95% confidence interval to estimate the true percentage of the residents who shop monthly at the local mall.

(A) (0.355, 0.445)  
(B) (0.366, 0.434)  
(C) (0.377, 0.423)  
(D) (0.380, 0.420)  
(E) It cannot be determined from the information given

Answer

21. To use the two sample  $t$  procedure to perform a significance test on the difference between two means, we assume

I. The population standard deviations are known  
II. The samples from each population are independent  
III. The samples are taken from the same population

(A) I only  
(B) II only  
(C) III only  
(D) I and II  
(E) I and III

Answer

22. A data set's five-number summary is given below:

minimum = 22  
first quartile = 31.2  
median = 44.5  
third quartile = 59.8  
maximum = 67

There are no outliers in this data set.

Which of the following conclusions can be drawn from the data?

- I. The mean is less than 44.5.
- II. Approximately 25% of the scores are above 59.8.
- III. Approximately 50% of the scores lie between 31.2 and 59.8.

- (A) I only
- (B) II only
- (C) II only
- (D) II and III only
- (E) I, II, and III

Answer

23. Which of the following statements related to residuals are true?

- I. The mean of the least-squares residuals is always zero.
- II. If one tries to fit a linear model to bivariate data, a curved pattern in a residual plot shows that the relationship between two variables is not linear.
- III. A residual plot can be a scatter plot of the regression residuals against the explanatory variable.

- (A) I only
- (B) II only
- (C) III only
- (D) II and III only
- (E) I, II, and III

Answer

24. Which of the following statements is true concerning a normal distribution curve?

- I. The curve is symmetric about its mean.
- II. The curve is always symmetric about 0.
- III. The area under the curve lying within one standard deviation of the mean contains approximately 50% of the scores.

- (A) I only
- (B) II only
- (C) III only
- (D) I and III only
- (E) I, II, and III

Answer

25. Which one of the following is NOT a continuous variable?

- (A) shoe sizes for women
- (B) depth of lakes in the Finger Lakes region
- (C) the speed of fastballs thrown by New York Yankees pitchers as measured by a radar gun
- (D) body temperatures of a group of aerobic participants after completing a workout
- (E) the number of minutes it takes for students to complete a test

Answer

26. You want to compute a 90% confidence interval for the mean of a population with unknown population standard deviation. The sample size is 30. The value of  $t$  you would use for this interval is

- (A) 0.90
- (B) 1.311
- (C) 1.645
- (D) 1.699
- (E) 1.96

Answer

27. A researcher wishes to test the following hypothesis

$$H_o: \mu = 100$$

$$H_a: \mu < 100$$

If the population mean is actually 95, but the researcher concludes that the mean is 100, what type of error has occurred?

- (A) Type I error
- (B) Type II error
- (C) Standard error of the mean
- (D) No error has occurred
- (E) The data does not provide enough information to determine the type of error

Answer

28. A set of data has the following five-number summary:

minimum = 17, first quartile = 27,  
median, = 40, third quartile = 49, maximum = 90

Which of the sets of numbers below contain ALL outliers?

- (A) 75, 80, 85
- (B) 78, 80, 85, 90
- (C) 83, 85, 90
- (D) 2, 3, 85, 90
- (E) 0, 80, 84, 89

Answer

29. When performing a test of hypothesis, the decision to choose either a  $z$ -test or a  $t$ -test is determined by the answers to which of the following questions?
- I. Is the sample mean greater than 30?
  - II. Is the standard deviation  $\sigma$  of the population known?
  - III. What is the level of significance  $\alpha$  for this test?
- (A) I only
  - (B) II only
  - (C) III only
  - (D) I and II
  - (E) II and III

Answer

30. A consumer group believes that boxes of Sugar Toasties Cereal contain, on average, less than the 15 ounces of cereal as advertised. They randomly sample 36 boxes of this cereal and find the mean  $\bar{x}$  is 14.9 and standard deviation  $s = 0.43$ . Assume that the weights are normally distributed and that  $\alpha = 0.05$ . In testing the consumer group's claim, what is the  $p$ -value for this hypothesis test?
- (A)  $-0.6511$
  - (B)  $0.0001$
  - (C)  $0.0500$
  - (D)  $0.0859$
  - (E)  $0.2547$

Answer

31. Which of the following are characteristics of a sound experimental design?
- I. Compare several treatments to control the effects of lurking variables on the response.
  - II. Use randomization to assign subjects to treatments.
  - III. To reduce chance variation in the results, repeat the experiment on many subjects.
- (A) I only  
(B) II only  
(C) III only  
(D) I and III only  
(E) I, II, and III

Answer

32. A newspaper in New Jersey conducted a poll of its readers to obtain their opinions on whether Michael Jordan would return to play basketball again in the NBA. 40% of the eligible readers participated. This is an example of a(n)
- (A) census  
(B) sample survey  
(C) experiment  
(D) observational study  
(E) systematic random sample

Answer

33. Which of the following is a resistant measure?
- (A) mean  $\mu$   
(B) median  
(C) standard deviation  $\sigma$   
(D) correlation coefficient  $r$  between two variables  $x$  and  $y$   
(E) range

Answer

34. Creating a sample of students by first dividing the entire student body into four groups: freshmen, sophomores, juniors, and seniors, and then selecting a simple random sample (SRS) from each of the four groups is an example of

(A) random sampling  
(B) cluster sampling  
(C) stratified sampling  
(D) systematic sampling  
(E) convenience sampling

Answer

35. The grades on a statistics exam were raised five points on each student's paper. This adjustment in the students' grades resulted in

I. increasing the class mean by five points.  
II. increasing the class median by five points.  
III. increasing the standard deviation by five points.

(A) I only  
(B) II only  
(C) III only  
(D) I and II only  
(E) I, II, and III

Answer

36. The equation of the least squares regression line for a set of points in a scatterplot is given by  $\hat{y} = 1.3 + 0.27x$ . The point (3, 2) is one point on this scatterplot. Which of the following is the residual for the point (3, 2)?

(A) -0.11  
(B) 0.11  
(C) 0.22  
(D) 1.04  
(E) 1.57

Answer

37. On a science examination given to 100 students, 50 earned scores of 90. Most of the other students scored 80, and the remaining students scored 20. Which of the following statements is true about the distribution of the scores?

- (A) The mode is equal to the mean.
- (B) The median is greater than the mean.
- (C) The mode is less than the mean.
- (D) The mean is greater than the median.
- (E) The mean is equal to the median.

Answer

38. Which statement(s) below concerning two normal distribution curves are necessarily true?

- I. Two normal curves can have the same mean, but different standard deviations.
- II. Two normal curves can have the same standard deviation, but different means.
- III. For two normal curves, the curve with the larger mean must also have the larger standard deviation.

- (A) I only
- (B) II only
- (C) III only
- (D) I and II only
- (E) I, II, and III

Answer



39. A test of hypothesis is used to compare two population proportions. Sample data drawn from each of the two populations is shown below where  $n$  is the sample size and  $x$  is the number of successes.

$$\begin{array}{ll} n_1 = 748 & n_2 = 614 \\ x_1 = 220 & x_2 = 192 \end{array}$$

The pooled sample proportion for this hypothesis test is

- (A) 0.2941
- (B) 0.3025
- (C) 0.3127
- (D) 0.4575
- (E) 0.6068

Answer

40. Which of the statements is true ?

- I. Two variables are confounding if their effects on a response variable cannot be distinguished from each other.
- II. A lurking variable has an effect on the relationship among variables in a study, but is not included among the variables studied.
- III. Observational studies of the effect of one variable on another variable can fail if the explanatory variable is confounded with a lurking variable.

- (A) I only
- (B) II only
- (C) III only
- (D) I and II only
- (E) I, II, and III

Answer